

<https://helda.helsinki.fi>

Genetic assessment reveals no population substructure and divergent regional and sex-specific histories in the Chachapoyas from northeast Peru

Guevara Torres, Evelyn

2020-12-31

Guevara Torres , E , Palo , J U , Översti , S , King , J L , Seidel , M , Stoljarova , M , Wendt , F R , Bus , M M , Guengerich , A , Church , W B , Guillen , S , Roewer , L , Budowle , B & Sajantila , A 2020 , ' Genetic assessment reveals no population substructure and divergent regional and sex-specific histories in the Chachapoyas from northeast Peru ' , PLoS One , vol. 15 , no. 12 , e0244497 . <https://doi.org/10.1371/journal.pone.0244497>

<http://hdl.handle.net/10138/325375>

<https://doi.org/10.1371/journal.pone.0244497>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

RESEARCH ARTICLE

Genetic assessment reveals no population substructure and divergent regional and sex-specific histories in the Chachapoyas from northeast Peru

Evelyn K. Guevara^{1*}, Jukka U. Palo^{1,2}, Sanni Översti³, Jonathan L. King⁴, Maria Seidel⁵, Monika Stoljarova⁶, Frank R. Wendt^{4,7,8}, Magdalena M. Bus^{4,8}, Anna Guengerich⁹, Warren B. Church¹⁰, Sonia Guillén¹¹, Lutz Roewer¹⁰, Bruce Budowle^{4,8}, Antti Sajantila^{1,12*}



1 Department of Forensic Medicine, University of Helsinki, Helsinki, Finland, **2** Forensic Genetics Unit, Finnish Institute for Health and Welfare, Helsinki, Finland, **3** Department of Biosciences, University of Helsinki, Helsinki, Finland, **4** Center for Human Identification, University of North Texas Health Science Center, Fort Worth, Texas, United States of America, **5** Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité – Universitätsmedizin Berlin, Berlin, Germany, **6** Department of Chemistry and Biotechnology, Tallinn University of Technology, Tallinn, Estonia, **7** Department of Psychiatry, Yale University School of Medicine and VA Connecticut Healthcare System, West Haven, Connecticut, United States of America, **8** Department of Microbiology, Immunology and Genetics, Graduate School of Biomedical Sciences, University of North Texas Health Science Center, Fort Worth, Texas, United States of America, **9** Eckerd College, Saint Petersburg, Florida, United States of America, **10** Department of Earth and Space Sciences, Columbus State University, Columbus, Georgia, United States of America, **11** Centro Mallqui, San Isidro, Lima, Peru, **12** Forensic Medicine Unit, Finnish Institute for Health and Welfare, Helsinki, Finland

* evelyn.guevara@helsinki.fi (EKG); antti.sajantila@helsinki.fi (AS)

OPEN ACCESS

Citation: Guevara EK, Palo JU, Översti S, King JL, Seidel M, Stoljarova M, et al. (2020) Genetic assessment reveals no population substructure and divergent regional and sex-specific histories in the Chachapoyas from northeast Peru. PLoS ONE 15(12): e0244497. <https://doi.org/10.1371/journal.pone.0244497>

Editor: Badri Padhukasahasram, Illumina Inc, UNITED STATES

Received: May 27, 2020

Accepted: December 10, 2020

Published: December 31, 2020

Copyright: © 2020 Guevara et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: A.S.: Medicinska Understödsföreningen Liv och Hälsa (no grant number) <http://www.livochhalsa.fi> The Finnish Society of Sciences and Letters (no grant number) <https://scientiarum.fi/eng/> E.K.G.: Kone Foundation (no grant number) <https://koneensaatio.fi> Regional Government of Amazonas, Peru (no grant number) Municipality of

Abstract

Many native populations in South America have been severely impacted by two relatively recent historical events, the Inca and the Spanish conquest. However decisive these disruptive events may have been, the populations and their gene pools have been shaped markedly also by the history prior to the conquests. This study focuses mainly on the Chachapoya peoples that inhabit the montane forests on the eastern slopes of the northern Peruvian Andes, but also includes three distinct neighboring populations (the Jívaro, the Huancas and the Cajamarca). By assessing mitochondrial, Y-chromosomal and autosomal diversity in the region, we explore questions that have emerged from archaeological and historical studies of the regional culture (s). These studies have shown, among others, that Chachapoyas was a crossroads for Coast-Andes-Amazon interactions since very early times. In this study, we examine the following questions: 1) was there pre-Hispanic genetic population substructure in the Chachapoyas sample? 2) did the Spanish conquest cause a more severe population decline on Chachapoyan males than on females? 3) can we detect different patterns of European gene flow in the Chachapoyas region? and, 4) did the demographic history in the Chachapoyas resemble the one from the Andean area? Despite cultural differences within the Chachapoyas region as shown by archaeological and ethnohistorical research, genetic markers show no significant evidence for past or current population substructure, although an Amazonian gene flow dynamic in the northern part of

Chachapoyas (no grant number) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

this territory is suggested. The data also indicates a bottleneck c. 25 generations ago that was more severe among males than females, as well as divergent population histories for populations in the Andean and Amazonian regions. In line with previous studies, we observe high genetic diversity in the Chachapoyas, despite the documented dramatic population declines. The diverse topography and great biodiversity of the northeastern Peruvian montane forests are potential contributing agents in shaping and maintaining the high genetic diversity in the Chachapoyas region.

Introduction

Chachapoya culture developed in a singular and biodiverse zone between the Andean mountain range and the Amazonian rainforests from northeastern Peru, and flourished around 900 CE (Common Era). Populations thrived in the area until conquered by the Incas (around 1475 CE) and Spaniards (first half of 16th Century). Since the Chachapoyas region has been an important area even before the Late Intermediate Period (1000–1475 CE), it has been a target of various archaeological, ethnohistoric, and genetic studies. Previous genetic assessments have demonstrated that the Chachapoyas region today harbors, unlike many other regions inhabited by Native South Americans, high levels of genetic diversity [1, 2], and they also have succeeded in placing Chachapoya populations on the genetic map of South America. However, several questions arising from the fields of archaeology and ethnohistory, remain to be explored with genetic data. These focus on the period of Inca and Spanish conquests in the 15th and 16th century and include a long-standing question regarding population substructure within Chachapoyas, as well as the sex-specific demographic impacts of the conquests. At a somewhat broader scale, we consider the issue of Chachapoyan affinities to human groups in the highland Andes and in lowland Amazonia.

Internal population substructure is of interest as the Chachapoya most likely have not been a homogeneous culture or single ethnic group. In fact, the name “Chachapoyas” was originally given by the Incas to denote an administrative province that held an aggregate of ethnic groups, and the name was later retained by the Spanish [3–5]. Similarly, in archaeology, Chachapoyas is still used as a shorthand to refer collectively to the several culturally diverse groups that specialists distinguish as inhabitants of this region before the Inca and Spanish conquests [6]. It is however unclear, whether the cultural differences identified also signify restricted gene flow and genetic segregation of these subgroups.

As in many other parts of South America, the Incan and European conquests had a profound demographic impact on the human groups in the Chachapoyas region, especially on males, who were particularly recruited for state-labor as well as warfare [4, 6, 7]. Around the years 1450–1475 CE, the Chachapoya territory was annexed as a province within the vast Inca realm, called Tawantinsuyo. Ethnohistorical studies indicate that after the Inca conquest, Chachapoya people were transferred to 18 different locations across Tawantinsuyo [4, 8, 9] and various groups of migrants were relocated into the Chachapoyas province from areas including Cajamarca, the North Coast, and Xauxa [4]. During the Inca civil war thousands of Chachapoya troops were recruited by Huascar to fight against Atahualpa [4]. Later, during the early colonial period, Chachapoyan males were also conscripted into the Spanish militia to fight against the Incas [10] and to explore unconquered Amazonian rainforest territories [4, 10]. European settlers also introduced pathogens that caused a series of recurrent epidemics that decimated the indigenous populations across Peru [11]. In Chachapoyas region, this

population decline is reported as dramatic and comparable to the ones experienced in coastal cities, the ports of entry of European migrants [6, 11].

Following the Spanish conquest, the introduction of European genes created an asymmetric genomic landscape, with male-mediated gene flow [1] but also different regions within Chachapoyas were likely being differently impacted by this process. In light of this history, the high native genetic diversity in the Chachapoya region reported today [1, 2] is intriguing.

Here we address the demographic impact in view of the population history outlined above, using autosomal, mitochondrial and Y-chromosomal marker data produced for four north-eastern Peruvian groups: Chachapoya, Jívaro, Huancas and Cajamarca. We examine the following questions: 1) was there pre-Hispanic genetic population substructure in the Chachapoyas sample? 2) did the Spanish conquest cause a more severe population decline on Chachapoyan males than on females? 3) can we detect different patterns of European gene flow in the Chachapoyas region? and 4) did the demographic history in the Chachapoyas resemble the one from the Andean area?

Materials and methods

This study followed ethical guidelines and standards (Helsinki declaration) and was approved by Helsinki University Hospital Ethical Committee (Decision #329/13/03/00/13) as previously described in [1]. Local board permits from Peru were previously obtained (Decision letters: N° 140-2009-GRAMAZONAS-PR and N° 04-2010/GRAMAZONAS-PR). Written informed consents were obtained from all participants. Altogether 246 individuals from three regions were included in this study: 1) the historical area identified as Chachapoyas, including the enclave population from the town of Huancas, 2) two Jivaroan (Chicham) populations (Awajún and Wampís) from the Amazonian rainforest, and 3) individuals from the neighboring Andean region of Cajamarca (The location of study populations can be found on Fig 1 of reference #1). These samples were genotyped by sequencing entire mitochondrial genomes ($N = 172$), Y-chromosomal ($N = 190$) and autosomal markers ($N = 246$).

For the analyses of substructure within Chachapoyas, the samples were divided into seven subgroups: Pomacochas, Chillao, Corobamba, Rodríguez de Mendoza, Chachapoya, La Jalca and Leymebamba. The assignment of individuals into a given subgroup was based on the place of birth up to the grandparents' generation and followed the strategy outlined in [1]. Briefly, individuals from nearby towns and villages were pooled into the same subgroup if there was documented historical affiliation (colonial and/or modern sources) or if they were united by archaeologically distinct patterns of material culture.

Genotyping

DNA extractions and Y chromosome short tandem repeat (STR) genotyping have been described in [1]. Full mitochondrial genomes were sequenced following a protocol described in [12]. The VCF files obtained from the Illumina MiSeq FGx system were converted into haplotypes using MitoSAVE [13]. Mitochondrial variants were curated manually and verified independently by three authors (EKG, FW and MS) using Integrative Genomic Viewer (IGV) software [14]. Poly-C stretches at positions 303–315 and 16184–16193 were excluded from subsequent analyses. Sequences were aligned to the revised Cambridge Reference Sequence (rCRS) [15] using CLUSTAL Omega [16, 17].

For the Y single nucleotide polymorphisms (Y-SNPs) analyses eight multiplex assays were used: MI and MII [18]; SpecE, SpecI and SpecJ [19]; SpecR [20]; and SpecQ-M346 and SpecQ-M3 (S1 Table)

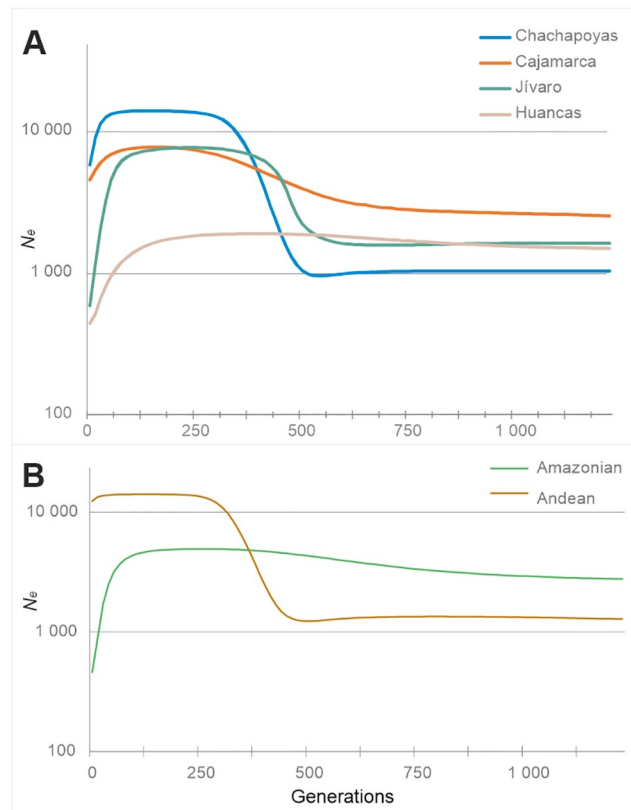


Fig 1. Coding region BSPs. (A) Study populations. (B) Macro-regions Andean ($N = 72$) and Amazonian ($N = 52$).

<https://doi.org/10.1371/journal.pone.0244497.g001>

Autosomal genotyping was carried out using the ForenSeq™ DNA Signature Prep Kit (Verogen; San Diego, Ca., USA) as described elsewhere [21]. Output FASTQ files from the ForenSeq™ Universal Analysis Software (UAS) pipeline were processed with STRait Razor v2s [22] to obtain the genotypes.

After data curation and filtering, 164 mitochondrial genomes (GenBank: MN857248–MN857411), 177 Y-chromosome (new Y-SNPs data) and 241 autosomal (27 short tandem repeats, STRs, and 94 single nucleotide polymorphisms, SNPs) profiles from four Peruvian populations (Chachapoyas, Huancas, Jívaro and Cajamarca) were used in downstream analyses.

Data analysis

Mitochondrial genomes. *Haplogroup composition, basic statistics and genetic distances.* For the analyses, the newly produced mitochondrial genomes were merged with data from [2]. Additionally, data from 78 worldwide reference populations were included (S2 Table).

Haplogroup assignments were obtained from HaploGrep v2.2 based on Phylotree 17 [23, 24]. Haplotypes belonging to non-Native American haplogroups H and U were excluded from the analyses addressing pre-contact population substructure and past effective population size. The effect of sample sizes on observed haplogroup frequencies was tested by permutations. Here, the number of each macro haplogroup observed was contrasted against the distribution obtained by 100,000 draws of N haplotypes, with N corresponding to the sample size in each

subpopulation. In the permutations, the success of drawing a certain haplotype was determined by its frequency in the modern Peruvian reference population ($N = 442$).

Summary statistics, e.g. haplotype diversity (h), nucleotide diversity (π), mean number of pairwise differences ($MNPD$), and pairwise genetic distances (Φ_{ST}) were estimated using Arlequin ver. 3.11 [25]. Genetic distances between populations were visualized with multi-dimensional scaling (MDS) plots built in RStudio version 3.5.3 with the package *cmdscale*.

Bayesian Skyline Plots (BSPs). Maternal demographic histories were explored with BSPs constructed using mitochondrial coding region data (nucleotides 577–16023) at three levels: 1) local (Chachapoyan subgroups), 2) regional (Chachapoyas, Huancas, Jivaro, Cajamarca), 3) subcontinental (Andean and Amazonian macro-regions). The best-fit substitution model for each population was determined with bModelTest [26]; the substitution model used for each population are described in S3 Table. BSPs were constructed using BEAST 2 v.2.4.7 [27] assuming discrete gamma distribution with four categories. As a tree prior, the piecewise-linear Skyline coalescent model with ten groups was used. The prior normal distribution for mutation rate was estimated based on six previously published mutation rates for mitochondrial coding region [28–33]. The mean mutation rate used was $\mu = 1.546 \times 10^{-8} \pm 3.675 \times 10^{-9}$ (SD) substitutions/pair/year. For each population, two molecular clocks were tested: strict and log normal relaxed molecular clock and the best-fit molecular clock was determined with Akaike information criterion method (AICM) implemented in Tracer [34]. The resulting clock models used for each population are presented in S3 Table. For each BEAST run, the Markov chain Monte Carlo (MCMC) length was set to 10,000,000 steps, with the first 10% of the steps discarded as burn-in and the chain sampled every 1,000 steps. Three independent runs were performed and combined with LogCombiner (part of BEAST package). Adequate effective sample sizes ($ESS > 200$) were checked in Tracer v.1.7.0 [35] for each of the parameters.

Y-STRs and Y-SNPs. *Haplogroup composition, basic statistics and genetic distances.* Similar to the mitochondrial DNA data, our Chachapoya Y-chromosomal dataset was merged with the data from [2]. In addition, STR data from 88 worldwide (23 STRs) and 79 Native American (17 STRs) reference populations were included (S2 Table). The Y-haplogroups were determined based on the Y-SNP data and also estimated based on the Y-STR haplotypes using a Haplogroup Predictor (<http://www.hprg.com/hapest5/>). Summary statistics were calculated in a similar fashion as the mitochondrial data using Arlequin ver.3.11.

Bayesian skyline plots. Demographic histories were assessed by constructing BSPs for the Chachapoyas pooled, as well as for one Andean and one Amazonian set. The male demographic history was explored by reconstructing Bayesian skyline plots with BEASTvntnr package [36] implemented in BEAST v2.4.7. The Sainudiin model [37] with computed frequencies was used, together with the uncorrelated relaxed lognormal clock model. Sample sizes, other prior distributions and parameters are displayed in S4 Table. Tree calibration was accomplished by setting the prior distribution for mutation and the node calibration, and by assuming an evolutionary mutation rate of $8.2 \times 10^{-4} \pm 5.7 \times 10^{-4}$ mutations/locus/generation calculated based on loci specific estimates presented in [38]. A node calibration approach was also used. However, as all Peruvian samples included belong to haplogroup Q its node calibration cannot be used directly. For this reason, an outgroup of six individuals from haplogroup R (sub-haplogroups R1a1a, R1a2, R1b11, R1b3, R2a and R2; extracted from [39]) were included in the analysis and the age estimate for a common node P1 was fixed. The P1 nodal age was set according to the divergence date determined based on the Y-chromosomal sequence data [39]. We assumed a normal distribution with mean = 1411 generations and SD = 100. This covered the P1 age obtained in [39] (mean = 35282 years before present (ybp) and 95% highest posterior density [33662, 36917] ybp), assuming a male generation time of 25

years. The settings for each BEAST run, the number of runs, and downstream procedures with LogCombiner and Tracer v.1.7.0, were identical to the ones used for mitochondrial data.

Autosomal STR and SNP data. In case of the autosomal data, the Chachapoyan samples were allocated to subgroups in a slightly different manner than with the lineage markers. The Kuelap in Chachapoya, Uchucmarca in Leymebamba and the original Jivaroan sample collection groups Awajún and Wampís were analyzed as separate units. Additionally, 31 reference populations were used (S2 Table). For five of these populations 121 markers (27 autosomal STRs and 94 SNPs) were available and for the remaining populations less than 19 autosomal STRs. Allele frequencies were calculated and tested for statistically significant departures ($p < 0.05$) from Hardy-Weinberg Equilibrium (HWE) and Linkage disequilibrium (LD) using Arlequin ver. 3.11.

Genetic distances and population structure. Pairwise genetic distances (F_{ST}) between populations were calculated in Arlequin ver. 3.11. Genetic substructure and admixture proportions were calculated at three different hierarchical levels: 1) locally within Chachapoyas, 2) intra-continently (Americas) and 3) inter-continently, using STRUCTURE v 2.3.4 [40–43]. For the intracontinental analysis, published reference data from 15 populations from the Americas were incorporated. Due to the limited availability of reference data, the analyses were based on 11 and 15 aSTR loci. For the intercontinental analysis, reference populations from African Americans (AFA), European descendants in the USA (CAU), Asian Americans (ASN) and Hispanics (HIS) from [44] were included. Downstream analyses used 27 autosomal STRs and 94 autosomal SNPs.

For the STRUCTURE analyses, the parameter lambda (λ) was estimated for our whole dataset and the resulting value ($\lambda = 0.5$) was fixed in all runs. The parameter *alpha* (α) was estimated independently for each population with an initial value of one. The admixture model chosen was correlated alleles and 8–12 iterations were calculated in each run. Convergence in the hyperparameters such as *alpha* and *F* was observed after c. 60,000 steps in both burn-in and run length. Number of ancestry components, *K*, was chosen based on two criteria, the estimated log probability of the data ($\ln P(X|K)$ [40] and delta *K* (ΔK) [45]. Cluster permutations and plotting of successive values of *K* from 1 to 7 were carried out with CLUMPAK main pipeline [46]. An alternative method to explore population substructure was Principal Coordinates Analyses (PCoA), for which Genalex add-in for Excel [47] and PAST 3.24 [48] were used. In addition to these methods, Analysis of Molecular Variance (AMOVA) for all marker systems was also performed in Arlequin ver. 3.11 [25].

Results

Mitochondrial DNA

Haplogroup composition. Among the newly produced data three mitochondrial lineages previously unreported in Phylotree17 were observed. Roughly 97% of the samples belonged to Native American haplogroups A2, B2, C1 and D (Table 1). The non-Native American haplogroups H and U were only observed in the Chachapoyas set with a frequency of 4.9%. The most frequent haplogroup is B2 (43.9%), followed by haplogroups C1 (24.4%), D (16.5%) and A2 (12.2%). In the whole mitochondrial dataset, larger variability of sub-haplogroups is observed within haplogroups A2 and B2, the former being the most diverse with five different sub-clades (S5 Table).

Among the seven Chachapoyan subgroups, differences in haplogroup distribution were apparent, e.g. Chillao lacks haplogroup D, whereas the Rodriguez de Mendoza lacks haplogroup A2 (Table 1). In order to assess whether these observations were the result of chance alone, random sampling procedures were simulated for each study population and in each

Table 1. Absolute and relative (%) mitochondrial haplogroup frequencies for all study populations*.

Population/Haplogroup	N	Native American				European	
		A2	B2	C1	D	H	U
Chachapoya	20	5 (25)	6 (30)	3 (15)	5 (25)	1 (5)	0
Chillao	10	3 (30)	3 (30)	3 (30)	0	1(10)	0
Corobamba	9	1 (11.1)	2 (22.2)	4 (44.4)	2 (22.2)	0	0
La Jalca	9	1 (11.1)	5 (55.5)	2 (22.2)	1 (11.1)	0	0
Leymebamba	14	1 (7.1)	4 (28.6)	4 (28.6)	5 (35.7)	0	0
Pomacochas	13	1 (7.7)	10 (76.9)	1 (7.7)	1 (7.7)	0	0
Rodríguez de Mendoza	27	0	8 (29.6)	8 (29.6)	8 (29.6)	2 (7.4)	1 (3.7)
Chachapoyas pooled	102	12 (11.8)	38 (37.3)	25 (24.5)	22 (21.6)	4 (3.9)	1 (1)
Jívaro	42	5 (11.9)	27 (64.3)	9 (21.4)	1 (2.4)	0	0
Huancas	4	1 (25.0)	2 (50)	0	1 (25)	0	0
Cajamarca	16	2 (12.5)	5 (31.3)	6 (37.5)	3 (18.8)	0	0
ALL COMBINED	164	20 (12.2)	72 (43.9)	40 (24.4)	27 (16.5)	4 (2.4)	1 (0.6)

* Figures only from our datasets.

<https://doi.org/10.1371/journal.pone.0244497.t001>

Chachapoyan subgroup. In case of most haplogroups, the number of observations fall inside a simulated distribution assuming average Peruvian haplogroup frequencies. However, both in Jívaro and Pomacochas, haplogroup B2 was significantly more common than expected, as was haplogroup H in Chillao and Rodríguez de Mendoza. Rodríguez de Mendoza lacked haplogroup A2 completely, when one would expect 2–9 observations in a sample of 27 individuals, and haplogroup D was underrepresented in the Jivaroan sample (S1 Fig).

Summary statistics and genetic distances. Considering the merged datasets (this study and Barbieri et al. 2017), high levels of genetic diversity (e.g. h , π) were observed in all study populations (S6 Table) even when the non-Native American haplotypes were excluded from the analysis. Genetic distances among the Chachapoyan subgroups revealed that Pomacochas differentiated significantly from all the others ($\Phi_{ST} \geq 0.1$, $p \leq 0.05$; S7 Table) and Rodríguez de Mendoza differentiated from Chachapoya, Chillao and Pomacochas subgroups. The observed genetic distances increased consistently (by 1–20 percentage points) when haplotypes belonging to non-Native American haplogroups were excluded. Since the sample sizes of all Chachapoyan subgroups are relatively small, real affinities cannot be differentiated from the effects of sampling error. In that sense, only subgroups showing significant differentiation, such as Pomacochas and Rodríguez de Mendoza, were kept independent; while the remaining were merged into a combined set labeled Chachapoyas (S8 Table).

The Chachapoyas pooled had genetic affinities (non-significant) with the Cajamarca but remained distant to the Jívaro and the Huancas. Genetic distances in the context of the Americas are summarized in an MDS plot where the clumped Chachapoyas set is positioned near the centroid. The Pomacochas falls in a group that includes the Amazonian Jívaro and populations from Ecuador. The Rodríguez de Mendoza set situates instead close to ancient Peruvian populations (S2 Fig).

Mitochondrial BSPs. The Bayesian Skyline plots based on the mitochondrial coding region are shown in Fig 1 and S3 Fig. The graphs exhibit a marked population growth in the distant past, which starts around 500 generations ago and reaches a plateau c. 250 generations ago. For the pooled Chachapoyas, as well as for all the subgroups, the plots suggest a reduction of N_e starting at around 125 generations ago (~3 kya) and persisting forward in time. The observed pattern resembles the one observed for the Andean region, but differs markedly from

the Amazonian one which do not show prominent growth at any stage in the past (Fig 1B). However, the reduction of N_e around 3 kya appears as a common feature in most study and reference populations.

Y-STRs and Y-SNPs

Haplogroup composition. Y-STR haplotypes were previously reported in [1]. Considering FTDNA order and equal priors, the new Y-SNP data revealed 12 differences in the haplogroup assignments by the STR haplogroup predictor used (S9 Table). Around 64% of the haplotypes belonged to Native American haplogroup Q followed by non-Native American haplogroup R (19.2%). The Chachapoyas pooled set and the Cajamarca have different proportions of non-Native American haplogroups other than R (Table 2). Locally, within Chachapoyas, haplogroup Q is found in high frequencies in all subgroups with the exception of Rodríguez de Mendoza where haplogroup R is predominant (41.2%). Overall, haplogroups Q, J, and R, each with three lineages, are the most diverse in the whole dataset (S10 Table).

Summary statistics and genetic distances. Considering the merged datasets (this study and Barbieri et al. 2017), all study populations exhibit high levels of diversity (e.g. h , π) when using Y-STR markers (S11 Table), which is also consistent when excluding the non-Native American haplotypes. Considering these 23 markers, genetic distances reveal that Pomacochas differentiated ($\Phi_{ST} \geq 0.1$, $p \leq 0.05$) from most of the other subgroups (S12 Table). After pooling the Chachapoyan subgroups, 17 STRs were used for comparisons in order to include more populations from the Americas (S12 Table). Here the Chachapoyas had affinities ($\Phi_{ST} \leq 0.05$, $p < 0.05$) only with the Huancas and remained distant to both the Jívaro and the Cajamarca. The Jivaroan set did not have any affinities with study or reference populations. In the MDS plot, the Chachapoyas pooled set falls in between Andean and Amazonian clusters very close to the Huancas. The subgroups Pomacochas and La Jalca are relatively close to each other but slightly distant to the pooled Chachapoyas. The Jívaro situates in the periphery of the plot, similar to other Amazonian groups, while the Cajamarca falls within an Andean cluster (S4 Fig).

Y-chromosomal BSPs. Considering the full Chachapoyas set and assuming all Y-STRs as one partition with $\mu = 8.2 \times 10^{-4} \pm 5.7 \times 10^{-4}$ per generation, a marked population growth is observed at around the same time as for the mitochondrial data, from 600–200 generations

Table 2. Absolute and relative (%) Y-chromosome haplogroup frequencies for all study populations*.

Population/Haplogroup	N	Native American	Eurasian					African
		Q	G-M201	I-P37.2	J	R	KL-M9	E-M35
Chachapoyas	24	12 (50.0)	1 (4.2)	1 (4.2)	1 (4.2)	8 (33.3)	0	1 (4.2)
Chillao	14	12 (85.7)	0	0	2 (14.3)	0	0	0
Corobamba	7	6 (85.7)	0	0	1 (14.3)	0	0	0
La Jalca	17	13 (76.5)	0	0	0	4 (23.5)	0	0
Leymebamba	19	8 (42.1)	0	1 (5.3)	5 (26.3)	3 (15.8)	0	2 (10.5)
Pomacochas	22	17 (77.3)	2 (9.1)	0	0	2 (9.1)	0	1 (4.5)
Rodríguez de Mendoza	17	2 (11.8)	3 (17.6)	1 (5.9)	3 (17.6)	7 (41.2)	1 (5.9)	0
Chachapoyas pooled	120	70 (58.3)	6 (5.0)	3 (2.5)	12 (10.0)	24 (20.0)	1 (0.8)	4 (3.3)
Jívaro	24	23 (95.8)	0	0	0	1 (4.2)	0	0
Huancas	12	11 (91.7)	0	0	0	1 (8.3)	0	0
Cajamarca	21	9 (43)	2 (9.5)	0	2 (9.5)	8 (38)	0	0
ALL COMBINED	177	113 (63.8)	8 (4.5)	3 (1.7)	14 (7.9)	34 (19.2)	1 (0.6)	4 (2.3)

*Figures only from our datasets.

<https://doi.org/10.1371/journal.pone.0244497.t002>

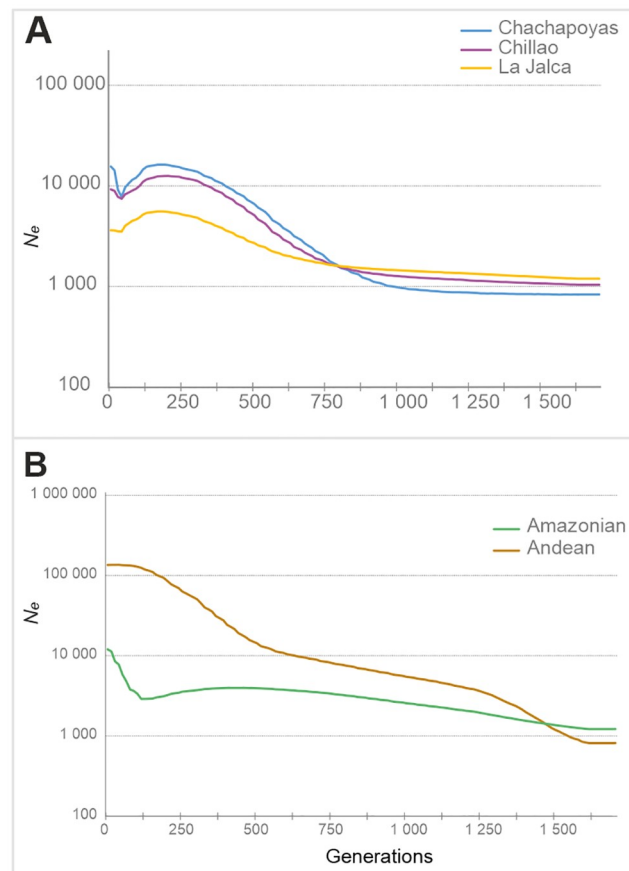


Fig 2. BSP with 21 Y-STRs treated as one partition. (A) For the Chillao, La Jalca and Chachapoyas pooled. (B) Macro-regions Andean ($N = 81$) and Amazonian ($N = 37$).

<https://doi.org/10.1371/journal.pone.0244497.g002>

(15–5 kya, assuming a male generation time of 25 years). Population decline for this set starts at around 200 generations and persists until it reaches its lowest point about 25 generations ago (c. 625 years ago assuming the mutation rate above) (Fig 2A). This deep N_e contraction of about 50% since the beginning of the population decline parallels either the time of Inca or Spanish conquest (22 or 25 years/generation interval, respectively). After this event, there is a continuous increase of the population size of more than 20% since reaching the lowest point, which is absent in the BSPs for mitochondrial data. Among the reference sets, the Amazonian one shows a very modest population increase starting at 800 generations ago (20 kya), which reaches a maximum height at 585 generations ago (~15 kya). After this, the population size declines slowly until reaching its lowest point at 110 generations ago (~2.8 kya). Following this recent population contraction, a rebound in N_e similar to the one observed in the Chachapoyas is detected for the Amazonian set (Fig 2B). The BSP for the Andean region in turn shows two intervals of population growth. The first one reaching its highest point at 512 generations ago (12.8 kya) and the second one at 73 generations ago (1.8 kya), after which there is no population decline in the recent past.

Autosomal

Summary statistics. All loci were in Hardy-Weinberg equilibrium in all populations with the exception of locus D1S1656 in the Chachapoyas. Linkage disequilibrium (LD), was

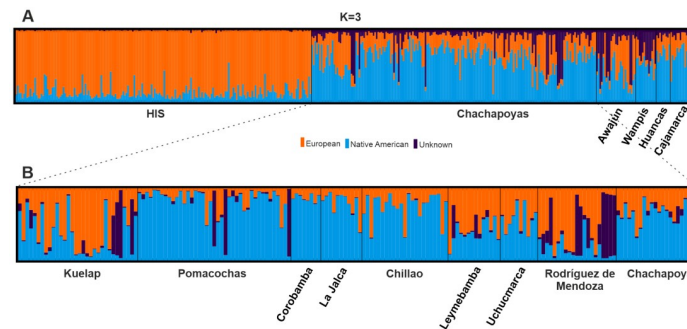


Fig 3. Admixture plots. (A) Study populations and the Hispanic reference (HIS) for $K = 3$. (B) Distribution of the 3rd unknown component within Chachapoyas. This component exhibits higher frequency in the Chachapoyan subgroups Kuelap, Pomacochas and Rodríguez de Mendoza.

<https://doi.org/10.1371/journal.pone.0244497.g003>

assessed for 27 aSTRs and 94 aSNPs in each study population. After correction for multiple testing ($p < 0.0001$), two pairs of loci showed departures from expectations, one in the Chachapoyas (rs13182883–rs338882), and one pair in the Cajamarca (rs6955448–rs917118). Genetic distances showed affinities ($F_{ST} < 0.05$, $p \leq 0.05$) among all study populations and the Hispanic set (HIS). The Caucasian dataset (CAU) in turn exhibited short genetic distances ($F_{ST} \leq 0.02$, $p \leq 0.05$) only to the Chachapoyas, the Awajún and the Cajamarca.

Population structure inference. Underlying substructure in the study populations was investigated hierarchically. At the upper level, $\ln P(X|K)$ and ΔK showed four ($K = 4$) well defined clusters represented by Africans (AFA), Europeans (CAU), Asians (ASN) and Native Americans (HIS and study populations) (S5 Fig). When only populations from the Americas were examined, the clustering showed two major components commonly found among most admixed populations from the Americas: European and Native American (Fig 3A). Additionally, the analysis showed a third unknown component present in much higher proportion in three study populations, Chachapoyas, Awajún and Wampís; but nearly absent in the reference Hispanic population (HIS). The 3rd unknown component was consistently observed with increasing number of K in plots including also other continental datasets (S1 File). Further examination, only within Chachapoyas (Fig 3B), showed consistently this 3rd component even when the analyses were restricted to a reduced number of markers (18–27 loci, not shown). This finding was supported by both $\ln P(X|K)$ and ΔK cross validation errors for $K = 3$ (S6 Fig). Zooming into the Chachapoyan subgroups, this third component found is more frequent in Kuelap, Pomacochas and Rodríguez de Mendoza (Fig 3B). The Native American component (light blue) is consistently observed at more than 50% in samples from Chachapoya, Chillao and Pomacochas subgroups. On the other hand, the European component (orange) is often found at higher frequencies ($> 50\%$) in Leymebamba, Rodríguez de Mendoza and Kuelap.

PCoA and genetic distances. When performing a PCoA including only the reference set having European ancestry (CAU) and the study populations (S7 Fig), it is possible to identify a small cluster composed of individuals only from three study populations, Chachapoyas, Awajún and Wampís. In this plot, the X-axis separates the European from the Native American clusters while the Y-axis separates the components within the Americas. The STRUCTURE output was used to better visualize the individuals that exhibit more than 60% of the third unidentified component. In order to accomplish this, all samples were binned by the proportion of the third unknown component (purple, in Fig 3), which resulted in three classes: light grey = 0–20%, grey = 21–59% and, black = 60–100% and then plotted in a new PCoA

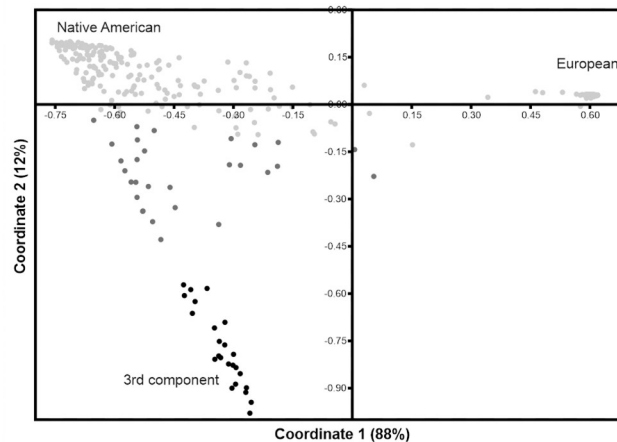


Fig 4. PCoA plot when samples are binned by their proportion of the 3rd unknown component. Individuals harboring more than 60% of this “3rd unknown component” in black.

<https://doi.org/10.1371/journal.pone.0244497.g004>

(Fig 4). The separation of European and Native American clusters is similar to the one described for S7 Fig. Individuals that harbor $\geq 60\%$ of the “3rd unknown component” fall close to each other (black dots) in a cline towards the Native American cluster.

To resolve the origins of the “3rd unknown component”, the genotypes of individuals harboring more than 60% of this component were placed in a new synthetic “population” and compared to other populations worldwide. These comparisons were carried out both with a larger set of markers 116 loci (S8A Fig) and with 10 aSTR loci (S8B Fig). With the 10 aSTR set, the “3rd unknown component” only shows affinities ($\Phi_{ST} \leq 0.05$, $p \leq 0.05$) with the Hispanic reference population from the US (HIS), urban populations from Mexico (Tijuana and Sonora), Chile (Calchaquí Valley) and a Mestizo sample from Nicaragua. On the other hand, the “3rd unknown component” remained distant ($\Phi_{ST} \geq 0.07$, $p \leq 0.05$) to the reference European populations, CAU from the US and Huelva from Spain. In the MDS plot with the full set of markers (116 loci), the “3rd unknown component” remains distant to all study and reference populations. With a reduced set of markers (10 aSTR loci) and more reference populations from Peru, the “3rd unknown component” falls in the periphery of the plot close to the Amazonian Cashibo (S8B Fig).

AMOVA was also used to investigate the distribution of diversity in a number of alternative subpopulation grouping schemes, based on the observed genetic distances between subgroups. As with the clustering analyses (STRUCTURE), AMOVA for autosomal data did not uncover population substructure within Chachapoyas, both F_{CT} (among groups) and F_{SC} (among populations within groups) values were non-significant and rather very small (S13 Table, run: 11_autos.arp). Here, more than 99.5% of the variation resided within subpopulations, regardless of the grouping.

In case of mitochondrial and Y-chromosomal data, the clusterings maximizing F_{CT} and minimizing F_{SC} values mirrored the patterns suggested by Φ_{ST} values (haplotypic AMOVA). The mtDNA distances indicate a grouping placing the Jívaro and the Pomacochas together and the remaining Chachapoya subsets in a different group, with the exception of Rodríguez de Mendoza which remains separate like the Huancas and the Cajamarca. For this scheme, 7.21% of the total variation resides between groups ($p < 0.001$) and 1.68% among populations within groups ($p < 0.05$) (S13 Table, run: 13_mito.arp). Similarly, for Y-chromosome data, the genetic distances suggested grouping only the Chachapoya subsets, Chillao, Chachapoya,

Rodríguez de Mendoza and Leymebamba but keeping the rest of the datasets separate. Following this, the AMOVA run 13_Y.arp, gave the best results for F_{CT} (0.104, $p < 0.003$) and F_{SC} (0.016, $p > 0.05$) values (S13 Table).

Discussion

The Chachapoya peoples flourished in the northeastern Peruvian montane forests until their encounter with the Inca and Spanish led gradually to the dissolution of the culture. Here we have assessed past population signals of the Chachapoyas persisting in the modern Peruvian genetic diversity. Despite cultural diversity among the Chachapoyas, we could not find clear signs of genetic substructure in this region. However, the high native genetic diversity observed sheds light on the complex, gender and region-specific population history of the Chachapoyas. Obviously, these inferences are to some extent limited by the contemporary data used, as the modern diversity has been shaped by the entire history of the target populations. Ancient DNA data could provide additional information, but only if sample sets representative in space and time could be obtained. Furthermore, one could also argue that the autosomal markers used in the ForenSeq kit, geared for forensic purposes, are not ideal for detecting substructure. This could stem from the relatively limited number of loci as well as from ascertainment bias, selective inclusion of loci in the forensics-aimed kits that maximize the diversity and power for identification of individuals (globally). Larger, genome-wide data sets would obviously offer better resolution for substructure detection, but the impact of the latter, ascertainment bias, is difficult to quantify for our data. Previous studies [49] have shown that 13 STR loci can detect substructure at population level, but due to the ascertainment bias, the distances are deflated, i.e. that the forensic markers are indeed less sensitive in detecting the population substructure. Similarly, as among the Chachapoyan subgroups no significant substructure was observed for more sensitive Y-chromosomal and mitochondrial data, the impact of the ascertainment bias is most likely of quantitative nature.

Lack of genetic substructure within Chachapoyas

It is becoming increasingly clear through recent ethnohistorical and archaeological research that the Chachapoyas region has hosted considerable social, political, and cultural diversity before the Late Horizon or Inca period (1475–1532 CE) [6, 50]. Within Chachapoyas, places separated by few dozens of kilometers like Luya, Chilchos, and Rodríguez de Mendoza, have held major cultural differences e.g. in mortuary practices, pottery styles and agricultural systems [51–56].

Despite these cultural clues from the past, signs of genetic subdivision in the region are weak. Although genetic distances and allele (haplotype) frequencies point to genetic differentiation of some Chachapoya subgroups, Y-chromosomal, mitochondrial and autosomal markers all convey conflicting pictures, thus evincing for rather insubstantial substructure (HWE, genetic distances, cluster analyses and AMOVA).

Only one of the seven subgroups, Pomacochas, differed significantly from most others in both mitochondrial and Y-chromosome data ($\Phi_{ST} > 0.070$). On the other hand, admixture analyses based on autosomal markers showed again no detectable population substructure within Chachapoyas. Here, regardless of the hierarchical setting (intercontinental, intracontinental and local), the Chachapoya samples clustered together (Fig 3 and S5 Fig). However, despite the lack of significant substructure within the Chachapoyas region, a few statistically significant differences of mitochondrial haplogroup distributions were observed, e.g. lack of haplogroup A2 in Rodríguez de Mendoza. All these observations suggest that there are true differences in the frequencies of some haplogroups and certain degree of differentiation at

some localities which echoes in the cultural diversity discussed above. Recent gene flow obviously may have obscured the past differentiation. Yet, ancient substructure is known to leave persistent demographic signals especially in the haploid markers, as shown e.g. in Finland [57].

DNA suggests Pomacochas—Jívaro connections. Instead of substructure and disconnection, mitochondrial and autosomal data shows that the northern section of the Chachapoyas region, the Pomacochas area, share genetic affinity with Jivaroan populations.

There is lack of archaeological research on whether the territory around Pomacochas and further north (Jumbilla and Yambrasbamba) was part of the Chachapoyas realm but potential historical connections between these two areas cannot be completely ruled out. It has been suggested that the historical Jívaro and the ancient Chachapoya peoples shared some cultural features [55, 58], however, the available archaeological evidence is yet insufficient to sustain this assumption. Despite this, the Chachapoyas region has been identified as a crossroads, at the center of interregional exchange networks throughout millennia [59]. Indeed, even early lithic assemblages found in the Chachapoyas region resemble lithic industries from as far as the north Peruvian coast (Paiján) and highland Ecuador (El Inga) [59] which account for the uniqueness of this region. From the genetics point of view, two recent independent studies [60, 61] have identified east-west gene flow linking the north Peruvian coast and highlands with Amazonian groups to the east of the northern Andes. Interestingly, one of these studies has found gene flow between the Chachapoyas and populations from the Jivaroan ethnolinguistic stock such as the Awajún and Candoshi [61]. This evidence gives further support to our finding despite the lack of published historical works on ancient or recent migrations from Jivaroan populations towards the Chachapoya territory or vice versa. As such, this finding provides new ground for archaeological and ethnohistorical research investigating cultural interactions between these areas.

Chachapoyas harbour an autosomal variation component of undetermined origin.

Among the Chachapoyas and the Jívaro, admixture analyses revealed— in addition to Native American and Caucasian components—a third autosomal component that could not be readily associated to any modern source population. Within Chachapoyas, inter-subgroup differences in the commonness of this component was observed especially in Rodríguez de Mendoza, but nearly absent in e.g. Chillao, La Jalca and Uchucmarca. On the regional level, the component was only present among the Chachapoyas and the Jivaroan Awajún and Wampís. In order to understand the origin of this variation, individuals harboring $\geq 60\%$ of this third component were regrouped and genetic distances between this group and reference populations from the Americas were estimated. These comparisons, however, were unable to shed light on the origin: the synthetic “3rd component group” showed no obvious affinity to reference datasets of African, European or Asian origin, nor to the populations included from the Americas. In analyses with more reference populations but less data (10 aSTRs), the group remained distant to all the other populations except the Amazonian Cashibo from Peru. The exact origin of this variation component remains thus elusive; it could derive from a yet unsampled ancient or modern population, for example representing assemblages that have developed locally in semi-isolated populations with no distinctive cultural features. On this note, a recent study [62] of South American populations has shown the existence of two autosomal components of alleged Amazonian origin: 1) The first one, distributed along the eastern slopes of the Northern Andes in Ecuador and Colombia, and 2) the second one, observed mostly in populations from the Peruvian rainforest. Due to geographic proximity and the commonness of the 3rd unknown component among the Amazonian Jívaro, we could suggest ties to the second component, although this cannot be tested with our current resolution.

Asymmetrical European gene flow in the Chachapoyas region

As in several previous studies on South American populations [63, 64], we observed gender-biased European gene flow into Chachapoyas: 5% of the mitochondrial and 42% of the Y-chromosomal haplotypes belonged to haplogroups typical for Europeans (such as H and U in mitochondrial; R and I in Y-chromosome). A substantial proportion of non-Native American Y-chromosomal lineages were also observed in Cajamarca, but this was not the case in Jívaro and Huancas.

During the conquest, the Spanish formed close alliances with the Chachapoyas against the Inca, and established the city of Chachapoyas as capital of the region. As the center of Spanish rule, and currently the largest town in the region, the Chachapoya subgroup (which includes also samples from Levanto) would be expected to host a higher proportion of European alleles but the highest European contribution (autosomes $> 50\%$) was found instead in Kuelap, Leymebamba, and Rodríguez de Mendoza. In fact, the proportion of non-Native American haplotypes peaked in Leymebamba and Rodríguez de Mendoza ($\geq 58\%$). Of these three areas, high European contribution in Leymebamba could be linked to its role as a gatekeeper to the Chachapoyas region, since it is one of the most important settlements on the eastern side of the mountains along the only road in the region that crosses the Andes and the Marañón River. This road system connected the Chachapoya and the Cajamarca regions, as well as Uchumarca in the south, already before the Spanish conquest [4], which may also have facilitated non-Native American gene flow during the Early Colonial Period. In the area around Rodríguez de Mendoza, Spanish settlements were already observed during a visit of archbishop Mogrovejo to the area in 1593 [65]. Thus, this area became a preferred destination for the permanence of European settlers echoing also in the genetic makeup of local residents. This was probably the result of a combination of factors, such as warmer climate, fertile soil and large rivers. In addition, its location, closer to the lush Amazonian rainforests, may have been of interest for Spanish conquistadors when pursuing expeditions in search of El Dorado [4, 10] which may have prompted the settlement of more Spanish people in the area. The high European genetic contribution observed in the Kuelap group is, in turn, slightly puzzling as the Inca and Spanish cultural presence on the left bank of the Utcubamba river, where this sample set comes from, seems to have been less significant than on the other side [4, 51, 59, 66]. Interestingly, the Chilaos subgroup harbors less European admixture, complying somehow with this pattern described by archaeology and ethnohistory. Although Kuelap was abandoned shortly after the Spanish incursions in this area, the site is recognized as one of the most important archaeological settlements within the Chachapoyas region and its significance as a center of religious and political power for the local populations [67, 68] may have transcended far beyond the Inca times. However, the site is seldom cited in early documents and was rediscovered late during the 19th century [6, 67]. The mismatches observed between the genetic and cultural evidence point to increasingly complex patterns of European interaction with the local populations at even smaller scales.

Demographic trajectories

For mitochondrial DNA, coalescent-based BSPs are extensively used methods to make inferences of the changes in the past female population size [69]. However, since for Y-STRs data similar approaches are not equally well established, past male population dynamics are not routinely assessed in population genetic studies. A comprehensive estimation of the male effective population sizes would require Y-chromosomal sequence data which is seldom produced. Nevertheless, to be able to compare the signals in the male and female population histories, we assumed an experimental approach to construct BSPs with Y-STR markers. The temporal

congruence in the Y-chromosome and mtDNA-based BSP patterns gives credence for the validity of the Y-chromosomal results obtained.

Thus, demographic histories in northeast Peru were explored with BSPs, using both mitochondrial sequences and Y-chromosomal STR haplotypes among the study populations. While the effective population sizes reconstructed by the BSPs are largely conjectural, the overall trends are revealing. The BSP patterns obtained from our data differed both, between sexes (markers), and regions (Andean/Amazonian).

In general, all BSP plots showed an ancient population expansion period that began at 625–500 generations ago i.e. 15.6–12.5 kya (a 25-year generation interval is assumed for all datings, unless indicated otherwise), a range that encompasses the earliest evidence of human presence in the continent [70, 71]. The connection of this early expansion to any real population events other than the entry in the Americas is nebulous, but the later developments are intriguing: the graphs show no signs for a mid-Holocene demographic crisis evident in South American archaeological radiocarbon data and palaeoclimatic evidence [72]. Instead, the Y-chromosomal and mitochondrial data demonstrate consistent patterns of population decline later in time, starting at around 125 generations ago i.e. 3.1 kya. This decline appears to temporally roughly coincide with a period of decreased solar irradiation (“The Homeric minimum”) which has been associated to global climatic change, including heightened El Niño oscillation [73, 74], and to major cultural shifts in Europe [75]. Interestingly, pollen and charcoal records from sediments of Lake Pomacochas in the Chachapoyas region demonstrate substantial changes in landscape around 3 kya, pointing somewhat paradoxically to heightened human activity. Increase on charcoal and *Poaceae* pollen, decrease of arboreal pollen and appearance of maize in the sedimentary record around this time all point to the clearing of forests for maize cultivation [76].

Sex-specific differences. The effective population size trajectories differ between sexes for the last 125 generations i.e. 3 kya. In the recent past, BSPs based on mitochondrial sequences show a steep but a relatively continuous decline which becomes severe over time but does not show a clear low point of population contraction. In contrast, the Y-chromosome BSP for the Chachapoyas sample set registers a steep population decline that reaches the lowest point c. 25 generations ago. This bottleneck would match the Inca or the Spanish conquest if a generation interval of 23 or 20 years, respectively, is assumed, which seem at least marginally realistic. These generation interval estimates would date the start of the general decline at 125 generations ago (above) to c. 2 850 and 2 500 BP.

The BSPs also suggest that the Chachapoya male population was nearly halved sometime between the 15th–17th centuries. It is tempting to associate demographic events as drastic as this with the Inca and Spanish conquests. During the Late Horizon (1475–1532 CE), the relocation policies applied by the Incas are recognized as important contributing factors to the depopulation in the Chachapoya region but these are not described as severe as the demographic collapse during and after European contact (multiple infectious diseases and pandemics over a prolonged time, war, selective relocation policies and forced labor) [11, 77]. However, all these events should have affected both males and females with equal acuteness, which is in contrast with the results. In addition to the above-mentioned evidence for general demographic collapse, males especially were affected by the Inca and Spanish military conscriptions. For instance, Chachapoyan males were recruited to fight the Inca as well as to participate in the early colonial El Dorado expeditions. Several of these expeditions to the tropical Amazonian rainforests started in towns within the Chachapoya region [10]. This would suggest that the severe male-biased population contraction may be associated with post-European contact events.

Regional patterns. In addition to the sex-specific differences, there are notable dissimilarities in the shape of BSPs between Andean and Amazonian reference populations. In general, the Andean male and female populations show higher effective population sizes for the last 250 generations, and clear signs of population growth in the past. The Amazonian populations exhibit a starkly contrasting pattern with smaller overall sizes and very moderate or non-existent population growth. This contrasting Andes/Amazon demographic pattern could reflect different levels of isolation and resource availability. This kind of regional dynamics has earlier been suggested for South America, based on Y-chromosomal data [78], and has also been proposed to explain global linguistic diversity [79]. Within Chachapoyas, most of the subgroups show Andean-type demographic trajectories, except Pomacochas and Corobamba, with patterns resembling those of the Amazonian populations. This adds to the earlier evidence of Amazonian (Jivaroan) affinity of the Pomacochas subgroup.

The growing effective population sizes in the Andean populations and in the Chachapoyas, as well as the overall high diversity in the Chachapoyas region, differ markedly from other South American Native populations. Interestingly, *in silico* analyses of species richness in South America from Late Quaternary until present have identified this region in northeastern Peru and southwestern Ecuador as one of the hotspots of biodiversity, also confirmed by empirical observations [80]. Here, birth and persistence of biological diversity in this region is most likely driven by topography and climate [80]. These rich ecosystems then provide more resources for the local human communities and could have contributed to the observed genetic diversity by allowing larger and more stable population sizes through time [81–83]. This area between the Amazonian rainforest and Andean mountains hosts different kind of ecosystems over a relatively short geographic distance, which may have buffered the local populations—human and non-human—in times of environmental change [84].

Concluding remarks

To the extent of our statistical resolution, we did not detect population substructure in the Chachapoyas sample, which accounts for the notion that genetic subdivision does not necessarily accompany cultural differentiation. In other regions of South America, population substructure has also been found to be negligible in groups from the same and different language stock [85–87]. In line with previous research [1, 2], we found exceptionally high levels of local genetic diversity in the Chachapoyas for both lineage markers that together with the large effective population sizes indicate particular demographic processes in this region driven most likely by topography and climate. This is also consistent with previous studies suggesting similar areas, at the junction of montane and tropical rainforests along the eastern slopes of the northern Andes, are as well identified as hotspots of human genetic diversity [86, 88–91]. In addition to these findings, we detected gene flow between the Amazonian Jívaro and an area situated at the north of the Chachapoyas territory here represented by the Pomacochas subgroup (which included also the localities of Jumbilla and Yambrasbamba). Gene flow events between neighboring territories such as these are not surprising as long-distance genetic exchanges of varying magnitude at this latitude have been observed between the Andean and Amazonian domains [91–93], and even with the Pacific Coast [60, 94]. Additionally, an unknown autosomal component has been identified among the Chachapoyas and the Awajún and Wampís. Although the origin of this component could not be clarified, we speculate it may be of Native American origin as it remains distant to African, European and Asian populations.

European admixture impacted the indigenous genetic variation at various scales, from macro-regions (e.g. Amazon, Andes) to populations [95, 96]. In this study, both lineage and

autosomal markers point to asymmetrical (mtDNA and Y-chromosome) European gene flow in the Chachapoya territory. For autosomal data these signatures of European admixture were heightened in the subgroups Kuelap, Leymebamba and Rodríguez de Mendoza. The geographic position and cultural importance of these areas are highlighted as explanations to these patterns, showing that differences can be detected even at smaller spatial scales. As the dates of admixture between the European and Native American components cannot be estimated from the current data, we should consider also the possibility that more recent admixture events may have concealed the patterns established during the Early Colonial Period.

The female demographic signature of the Chachapoyas as a whole and the Jivaroan set resemble the one from the Andean area whereas the male population history in the Chachapoyas is more similar to the Amazonian one. Larger long-term effective population sizes have been previously characterized for the populations from the Andean area which contrasts with other regions of South America [78, 97]. Moreover, comparisons of both male and female demographic histories also indicate a more severe population decline on males which finds support in the historical portrait of the northeast Peruvian Andes as an area that experienced a dramatic population decline, particularly among males [4]. Even within the Amazonian region, genome wide studies in South America have shown that population sizes have been shaped by different demographic processes throughout their history [62, 92]. In that sense, in addition to the detection of genetic drift by means of autosomal data, it is also vital to examine demographic histories based on lineage markers, as different processes affecting the male and female population sizes can result in divergent patterns like the ones observed in this and other studies e.g. [98]. We acknowledge that although various inferences about the population structure and population history in the Chachapoyas have been provided here, other questions such as the origin of the Chachapoya ancestry can ultimately be addressed with higher resolution and/or ancient DNA data.

Supporting information

S1 Fig. Permuted haplogroup distributions and actual observations showing the deviations ($p < 0.05$) with arrows in Chillao, Pomacochas and Jívaro.

(TIF)

S2 Fig. Mitochondrial MDS plot with populations from the Americas showing the Pomacochas and the Rodríguez de Mendoza subgroups as well as the pooled remaining subgroups under the label Chachapoyas.

(TIF)

S3 Fig. Mitochondrial coding region BSPs for all Chachapoyan subgroups.

(TIF)

S4 Fig. Y-chromosome MDS plot with populations from the Americas showing the Chachapoyan subgroups pooled with exception of La Jalca and Pomacochas.

(TIF)

S5 Fig. Admixture plot including study and reference populations for $K = 4$.

(TIF)

S6 Fig. In the Chachapoyas set, a value of $K = 3$ was consistently estimated by two methods, $\ln P(X|K)$ and ΔK .

(TIF)

S7 Fig. PCoA plot showing the study populations and the set CAU (US Americans with European ancestry). The circle indicates Individuals harboring > 60% of the “3rd unknown component”.

(TIF)

S8 Fig. MDS plots of autosomal genetic distances showing the individuals in the “3rd unknown component”. Study populations and Chachapoyan subgroups in black, m: mestizo. (A) 116 loci. (B) 10 loci.

(TIF)

S1 Table. Multiplexes M3 and M346. (A) Genotyping. (B) Primer information grouped by multiplexes.

(XLSX)

S2 Table. Sample sizes and reference populations. (A) mitochondrial DNA. (B) Y-Chromosome. (C) autosomal data.

(XLSX)

S3 Table. Sample sizes and best substitution models for mitochondrial DNA data BSPs.

(XLSX)

S4 Table. Sample sizes, supported clock model and prior parameters for Y-chromosome data BSPs.

(XLSX)

S5 Table. Absolute frequency of mitochondrial sub-haplogroups in all study populations. Frequencies only from our datasets.

(XLSX)

S6 Table. Diversity indices for mitochondrial DNA data. (A) Non-Native American haplotypes are included in the study populations. (B) Non-Native American haplotypes are excluded in the study populations.

(XLSX)

S7 Table. Mitochondrial Φ_{ST} distances. (A) Non-Native American haplotypes are included in the study populations. (B) Non-Native American haplotypes are excluded in the study populations.

(XLSX)

S8 Table. Mitochondrial Φ_{ST} distances when the Chachapoyan subgroups are pooled. Non-Native American haplotypes are excluded in the study populations.

(XLSX)

S9 Table. Y-chromosome haplotypes (Guevara et al. 2016) and SNP assignment by multiplexes (This study).

(XLSX)

S10 Table. Absolute frequency of Y-chromosome sub-haplogroups in all study populations. Frequencies only from our datasets.

(XLSX)

S11 Table. Diversity indices. (A) 23 Y-STRs. (B) 17 Y-STRs.

(XLSX)

S12 Table. Y-chromosome Φ_{ST} distances. (A) 23 Y-STRs when the Chachapoyan subgroups are not pooled. (B) 17 Y-STRs when the Chachapoyan subgroups are pooled.
(XLSX)

S13 Table. AMOVA results for different kind of marker systems.
(XLSX)

S1 File. STRUCTURE plots including study populations and datasets from different continental origins ($K = 2-6$).
(PDF)

Acknowledgments

We are grateful to all sample donors from the administrative region of Amazonas in Peru. We thank Xiangpei Zeng for his guidance during laboratory procedures. We thank Jouni Kärkkäinen for setting up a graphic user interface (GUI) to run STRUCTURE remotely and Elina Salmela for writing an R-script and running the calculations for haplogroup frequency permutations.

Author Contributions

Conceptualization: Evelyn K. Guevara, Jukka U. Palo, Antti Sajantila.

Data curation: Evelyn K. Guevara, Jonathan L. King.

Formal analysis: Evelyn K. Guevara, Sanni Översti.

Funding acquisition: Evelyn K. Guevara, Sonia Guillén, Antti Sajantila.

Investigation: Evelyn K. Guevara, Jonathan L. King, Maria Seidel, Monika Stoljarova, Frank R. Wendt, Magdalena M. Bus.

Methodology: Evelyn K. Guevara, Jukka U. Palo.

Project administration: Evelyn K. Guevara, Jukka U. Palo, Antti Sajantila.

Resources: Evelyn K. Guevara, Sonia Guillén, Lutz Roewer, Bruce Budowle, Antti Sajantila.

Supervision: Jukka U. Palo, Antti Sajantila.

Validation: Evelyn K. Guevara, Sanni Översti, Jonathan L. King, Maria Seidel.

Visualization: Evelyn K. Guevara, Sanni Översti.

Writing – original draft: Evelyn K. Guevara, Jukka U. Palo.

Writing – review & editing: Evelyn K. Guevara, Jukka U. Palo, Jonathan L. King, Monika Stoljarova, Frank R. Wendt, Anna Guengerich, Warren B. Church, Sonia Guillén, Lutz Roewer, Bruce Budowle, Antti Sajantila.

References

1. Guevara EK, Palo JU, Guillén S, Sajantila A. MtDNA and Y-chromosomal diversity in the Chachapoya, a population from the northeast Peruvian Andes-Amazon divide. *American Journal of Human Biology*. 2016:n/a-n/a. <https://doi.org/10.1002/ajhb.22878> PMID: 27265853
2. Barbieri C, Sandoval JR, Valqui J, Shimelman A, Ziemendorff S, Schroeder R, et al. Enclaves of genetic diversity resisted Inca impacts on population history. *Scientific Reports*. 2017; 7. <https://doi.org/10.1038/s41598-017-17728-w> PMID: 29234095
3. Schjellerup I. Reflexiones sobre los Chachapoya en el Chinchaysuyo. *Boletín de Arqueología PUCP*. 2002; 6:43–56.

4. Schjellerup I. Incas y españoles en la conquista de los chachapoya. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú and Instituto Francés de Estudios Andinos; 2005. 641 p.
5. Zevallos Q. J. El área geográfico-cultural de la prehistoria de Chachapoyas. Una nueva postulación. *Gaceta Arqueológica Andina*. 1995;24:13–23.
6. Church W, Guengerich A. Introducción: La reconstrucción de Chachapoyas a través de la historia e historiografía *Boletín de Arqueología PUCP*. 2017; 23:5–38.
7. Levine TY. Inka labor resvise at the regional level—The functional reality. *Ethnohistory*. 1987; 34(1):14–46.
8. Espinoza W. La etnia Chilcho y su enclave de mitmas en Cajamarca, siglos xv–xx. *Investigaciones Sociales*. 2006; X(16):193–238.
9. Espinoza W. *Etnohistoria Ecuatoriana: Estudios y Documentos*. Quito: Ediciones Abya-Yala; 1988.
10. Espinoza W. Los señoríos étnicos de Chachapoyas y la alianza hispano-chacha. Visitas, informaciones y memoriales inéditos 1572–1574. *Revista Histórica* 1967:224–332.
11. Cook ND. *Demographic Collapse: Indian Peru, 1520–1620*. Cambridge University Press; 1981.
12. Stoljarova M, King JL, Takahashi M, Aaspollu A, Budowle B. Whole mitochondrial genome genetic diversity in an Estonian population sample. *International Journal of Legal Medicine*. 2016; 130(1):67–71. <https://doi.org/10.1007/s00414-015-1249-4> PMID: 26289416
13. King JL, Sajantila A, Budowle B. mitoSAVE: Mitochondrial sequence analysis of variants in Excel. *Forensic Science International: Genetics*. 2014; 12:122–5. <https://doi.org/10.1016/j.fsigen.2014.05.013>.
14. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature Biotechnology*. 2011; 29:24. <https://doi.org/10.1038/nbt.1754> <https://www.nature.com/articles/nbt.1754#supplementary-information>. PMID: 21221095
15. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999; 23(2):147–. <https://doi.org/10.1038/13779> PMID: 10508508
16. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*. 2011; 7. <https://doi.org/10.1038/msb.2011.75> PMID: 21988835
17. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. *Protein Science*. 2018; 27(1):135–45. <https://doi.org/10.1002/pro.3290> PMID: 28884485
18. Geppert M, Roewer L. SNaPshot (R) Minisequencing Analysis of Multiple Ancestry-Informative Y-SNPs Using Capillary Electrophoresis. In: Alonso A, editor. *DNA Electrophoresis Protocols for Forensic Genetics*. Methods in Molecular Biology. 8302012. p. 127–40.
19. Nuñez C, Geppert M, Baeta M, Roewer L, Martinez-Jarreta B. Y chromosome haplogroup diversity in a Mestizo population of Nicaragua. *Forensic Science International-Genetics*. 2012; 6(6):E192–E5. <https://doi.org/10.1016/j.fsigen.2012.06.011> PMID: 22770600
20. Haas C, Shved N, Ruehli FJ, Papageorgopoulou C, Purps J, Geppert M, et al. Y-chromosomal analysis identifies the skeletal remains of Swiss national hero Jorg Jenatsch (1596–1639). *Forensic Science International-Genetics*. 2013; 7(6):610–7. <https://doi.org/10.1016/j.fsigen.2013.08.006> PMID: 24035510
21. Wendt FR, Churchill JD, Novroski NMM, King JL, Ng J, Oldt RF, et al. Genetic analysis of the Yavapai Native Americans from West-Central Arizona using the Illumina MiSeq FGx™ forensic genomics system. *Forensic Science International: Genetics*. 2016; 24:18–23. <https://doi.org/10.1016/j.fsigen.2016.05.008>.
22. King JL, Wendt FR, Sun J, Budowle B. STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. *Forensic Science International: Genetics*. 2017; 29:21–8. <https://doi.org/10.1016/j.fsigen.2017.03.013>.
23. Kloss-Brandstaetter A, Pacher D, Schoenherr S, Weissensteiner H, Binna R, Specht G, et al. Haplo-Grep: A Fast and Reliable Algorithm for Automatic Classification of Mitochondrial DNA Haplogroups. *Human Mutation*. 2011; 32(1):25–32. <https://doi.org/10.1002/humu.21382> PMID: 20960467
24. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation*. 2009; 30(2):E386–E94. <https://doi.org/10.1002/humu.20921> PMID: 18853457
25. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*. 2005; 1(EBO-1-Excoffier(Sc)):47–50.
26. Bouckaert RR, Drummond AJ. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *Bmc Evolutionary Biology*. 2017; 17. <https://doi.org/10.1186/s12862-017-0890-6> PMID: 28166715

27. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, et al. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology*. 2014; 10(4):e1003537. <https://doi.org/10.1371/journal.pcbi.1003537> PMID: 24722319
28. Ingman M, Kaessmann H, Pääbo S, Gyllenstein U. Mitochondrial genome variation and the origin of modern humans. *Nature*. 2000; 408(6813):708–13. <https://doi.org/10.1038/35047064> PMID: 11130070
29. Tang H, Siegmund DO, Shen PD, Oefner PJ, Feldman MW. Frequentist estimation of coalescence times from nucleotide sequence data using a tree-based partition. *Genetics*. 2002; 161(1):447–59. PMID: 12019257
30. Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, et al. Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100(1):171–6. <https://doi.org/10.1073/pnas.0136972100> PMID: 12509511
31. Soares P, Ermini L, Thomson N, Mormina M, Rito T, Roehli A, et al. Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *American Journal of Human Genetics*. 2009; 84(6):740–59. <https://doi.org/10.1016/j.ajhg.2009.05.001> PMID: 19500773
32. Ho SYW, Endicott P. The Crucial Role of Calibration in Molecular Date Estimates for the Peopling of the Americas. *The American Journal of Human Genetics*. 2008; 83(1):142–6. <https://doi.org/10.1016/j.ajhg.2008.06.014> PMID: 18606310
33. Fu Q, Mitnik A, Johnson Philip LF, Bos K, Lari M, Bollongino R, et al. A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. *Current Biology*. 2013; 23(7):553–9. <https://doi.org/10.1016/j.cub.2013.02.044> PMID: 23523248
34. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. Improving the Accuracy of Demographic and Molecular Clock Model Comparison While Accommodating Phylogenetic Uncertainty. *Molecular Biology and Evolution*. 2012; 29(9):2157–67. <https://doi.org/10.1093/molbev/mss084> PMID: 22403239
35. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*. 2018; 67(5):901–4. <https://doi.org/10.1093/sysbio/syy032> PMID: 29718447
36. Dhawan A. Bayesian inference of phylogeny using Variable Number of Tandem Repeats and Markov Chain Monte Carlo. Netherlands: Utrecht University; 2016.
37. Sainudiin R, Durrett RT, Aquadro CF, Nielsen R. Microsatellite mutation models: Insights from a comparison of humans and chimpanzees. *Genetics*. 2004; 168(1):383–95. <https://doi.org/10.1534/genetics.103.022665> PMID: 15454551
38. Shi W, Ayub Q, Vermeulen M, Shao R-g, Zuniga S, van der Gaag K, et al. A Worldwide Survey of Human Male Demographic History Based on Y-SNP and Y-STR Data from the HGDP-CEPH Populations. *Molecular Biology and Evolution*. 2010; 27(2):385–93. <https://doi.org/10.1093/molbev/msp243> PMID: 19822636
39. Karmin M, Saag L, Vicente M, Sayres MAW, Jarve M, Talas UG, et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research*. 2015; 25(4):459–66. <https://doi.org/10.1101/gr.186684.114> PMID: 25770088
40. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. 2000; 155(2):945–59.
41. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*. 2003; 164(4):1567–87. PMID: 12930761
42. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*. 2007; 7(4):574–8. <https://doi.org/10.1111/j.1471-8286.2007.01758.x> PMID: 18784791
43. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*. 2009; 9(5):1322–32. <https://doi.org/10.1111/j.1755-0998.2009.02591.x> PMID: 21564903
44. Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B. Characterization of genetic sequence variation of 58 STR loci in four major population groups. *Forensic Science International: Genetics*. 2016; 25:214–26. <https://doi.org/10.1016/j.fsigen.2016.09.007> PMID: 27697609
45. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. 2005; 14(8):2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> PMID: 15969739
46. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. 2015; 15(5):1179–91. <https://doi.org/10.1111/1755-0998.12387> PMID: 25684545

47. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 2012; 28(19):2537–9. <https://doi.org/10.1093/bioinformatics/bts460> PMID: 22820204
48. Hammer O, Harper DAT, Ryan PD. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*. 2001; 4(1):Unpaginated-Unpaginated.
49. Silva NM, Pereira L, Poloni ES, Currat M. Human Neutral Genetic Variation and Forensic STR Data. *Plos One*. 2012; 7(11). <https://doi.org/10.1371/journal.pone.0049666> PMID: 23185401
50. Ruiz A. Deslindes étnicos en la historia de Amazonas, Perú. *Boletín de Arqueología PUCP*. 2017; 23:41–56.
51. Crandall J. *The Transformations of Empire: Negotiating Inka and Spanish Colonialism in an Upper Amazonian Community*. Gainesville, FL: University of Florida; 2018.
52. Fabre O. Contribution à l'archéologie de la région Chachapoya, Pérou: Université Paris-Sorbonne; 2006.
53. Guengerich A. Domestic architecture and the materiality of public-making in pre-Columbian Eastern Peru. *Journal of Social Archaeology*. 2017; 17(3):263–84. <https://doi.org/10.1177/1469605317722217>
54. McCray B. Límites e interfaces de Chachapoyas en la región Rodríguez de Mendoza. *Boletín de Arqueología PUCP*. 2017; 23:187–206.
55. Koschmieder K. Los orígenes y el desarrollo de la organización socio-política de la cultura chachapoya: una mirada desde la provincia de Luya, Departamento Amazonas, Perú. In: Rostain S, editor. *Antes de Orellana-Actas del 3er encuentro internacional de arqueología amazónica*. Quito, Ecuador: Instituto Francés de Estudios Andinos; 2014. p. 243–9, 528–9.
56. Nystrom KC, Buikstra JE, Muscutt K. Chachapoya mortuary behavior: a consideration of method and meaning. *Chungara-Revista De Antropología Chilena*. 2010; 42(2):477–95. <https://doi.org/10.4067/s0717-73562010000200010>
57. Neuvonen AM, Putkonen M, Oversti S, Sundell T, Onkamo P, Sajantila A, et al. Vestiges of an Ancient Border in the Contemporary Genetic Diversity of North-Eastern Europe. *Plos One*. 2015; 10(7). <https://doi.org/10.1371/journal.pone.0130331> PMID: 26132657
58. Guffroy J. El Horizonte corrugado: correlaciones estilísticas y culturales. *Bulletin de l'Institut Français d'Études Andines*. 2006; 35(3):347–59.
59. Church WB, von Hagen A. Chachapoyas: Cultural Development at an Andean Cloud Forest Cross-roads. In: Silverman H, Isbell WH, editors. *The Handbook of South American Archaeology*. New York: Springer Science+Business Media, LLC; 2008. p. 906–23.
60. Nakatsuka N, Lazaridis I, Barbieri C, Skoglund P, Rohland N, Mallick S, et al. A Paleogenomic Reconstruction of the Deep Population History of the Andes. *Cell*. 2020; 181(5):1132–+. <https://doi.org/10.1016/j.cell.2020.04.015> PMID: 32386546
61. Borda V, Alvim I, Aquino MM, Silva C, Soares-Souza GB, Leal TP, et al. The genetic structure and adaptation of Andean highlanders and Amazonian dwellers is influenced by the interplay between geography and culture. *bioRxiv*. 2020:2020.01.30.916270. <https://doi.org/10.1101/2020.01.30.916270>
62. Barbieri C, Barquera R, Arias L, Sandoval JR, Acosta O, Zurita C, et al. The Current Genomic Landscape of Western South America: Andes, Amazonia, and Pacific Coast. *Molecular Biology and Evolution*. 2019; 36(12):2698–713. <https://doi.org/10.1093/molbev/msz174> PMID: 31350885
63. Bolnick DA, Bolnick DI, Smith DG. Asymmetric Male and Female Genetic Histories among Native Americans from Eastern North America. *Molecular Biology and Evolution*. 2006; 23(11):2161–74. <https://doi.org/10.1093/molbev/msl088> PMID: 16916941
64. Gayà-Vidal M, Moral P, Saenz-Ruales N, Gerbault P, Tonasso L, Villena M, et al. mtDNA and Y-chromosome diversity in Aymaras and Quechuas from Bolivia: Different stories and special genetic traits of the Andean Altiplano populations. *Am J Phys Anthropol*. 2011; 145(2):215–30. <https://doi.org/10.1002/ajpa.21487> PMID: 21469069
65. Mogrovejo TAd. Diario de la segunda visita pastoral que hizo de su arquidiócesis el Ilustrísimo Señor Don Toribio Alonso de Mogrovejo, arzobispo de los Reyes. *Revista del Archivo Nacional del Perú*. 1921 [1593]; 11(1):37–78.
66. Crandall J. El desarrollo espacial de las comunidades Chachapoyas bajo la dominación colonial Inka y Española. *BOLETIN DE ARQUEOLOGIA PUCP*. 2017; 23:283–311.
67. Narváez A. Kuélap: centro del poder político religioso de los Chachapoyas. *Los Chachapoyas*. Lima, Peru: Banco de Crédito; 2013. p. 87–159.
68. Toyne M, Narváez A. The Fall of Kuelap. *Bioarchaeological Analysis of Death and Destruction on the Eastern Slopes of the Andes*. In: Scherer AK, Verano JW, editors. *Embattled Bodies, Embattled Places War in Pre-Columbian Mesoamerica and the Andes*. Washington D.C.: Dumbarton Oaks Research Library and Collection; 2014. p. 341–64.

69. Ho SYW, Shapiro B. Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*. 2011; 11(3):423–34. <https://doi.org/10.1111/j.1755-0998.2011.02988.x> PMID: 21481200
70. Dillehay TD. Probing deeper into first American studies. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(4):971–8. <https://doi.org/10.1073/pnas.0808424106> PMID: 19164556
71. Faught MK. Archaeological roots of human diversity in the New World: a compilation of accurate and precise radiocarbon ages from earliest sites. *American Antiquity*. 2008; 73(4):670–98.
72. Riris P, Arroyo-Kalin M. Widespread population decline in South America correlates with mid-Holocene climate change. *Scientific Reports*. 2019; 9. <https://doi.org/10.1038/s41598-019-43086-w> PMID: 31073131
73. Sandweiss DH, Maasch KA, Burger RL, Richardson JB, Rollins HB, Amy C. Variation in Holocene El Niño frequencies: Climate records and cultural consequences in ancient Peru. *Geology*. 2001; 29(7):603–6. [https://doi.org/10.1130/0091-7613\(2001\)029<0603:VIHENO>2.0.CO;2](https://doi.org/10.1130/0091-7613(2001)029<0603:VIHENO>2.0.CO;2)
74. Wanner H, Beer J, Butikofer J, Crowley TJ, Cubasch U, Fluckiger J, et al. Mid- to Late Holocene climate change: an overview. *Quaternary Science Reviews*. 2008; 27(19–20):1791–828. <https://doi.org/10.1016/j.quascirev.2008.06.013>
75. Rach O, Engels S, Kahmen A, Brauer A, Martin-Puertas C, van Geel B, et al. Hydrological and ecological changes in western Europe between 3200 and 2000 years BP derived from lipid biomarker delta D values in lake Meerfelder Maar sediments. *Quaternary Science Reviews*. 2017; 172:44–54. <https://doi.org/10.1016/j.quascirev.2017.07.019>
76. Bush MB, Mosblech NAS, Church W. Climate change and the agricultural history of a mid-elevation Andean montane forest. *The Holocene*. 2015; 25(9):1522–32. <https://doi.org/10.1177/0959683615585837>
77. Koch A, Brierley C, Maslin MM, Lewis SL. Earth system impacts of the European arrival and Great Dying in the Americas after 1492. *Quaternary Science Reviews*. 2019; 207:13–36. <https://doi.org/10.1016/j.quascirev.2018.12.004>
78. Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, De Stefano GF, Labarga CM, et al. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. *American Journal of Human Genetics*. 2001; 68(6):1485–96. <https://doi.org/10.1086/320601> PMID: 11353402
79. Hua X, Greenhill SJ, Cardillo M, Schneemann H, Bromham L. The ecological drivers of variation in global language diversity. *Nature Communications*. 2019; 10. <https://doi.org/10.1038/s41467-019-09842-2> PMID: 31053716
80. Rangel TF, Edwards NR, Holden PB, Diniz JAF, Gosling WD, Coelho MTP, et al. Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science*. 2018; 361(6399). <https://doi.org/10.1126/science.aar5452> PMID: 30026200
81. Tallavaara M, Eronen JT, Luoto M. Productivity, biodiversity, and pathogens influence the global hunter-gatherer population density. *Proceedings of the National Academy of Sciences of the United States of America*. 2018; 115(6):1232–7. <https://doi.org/10.1073/pnas.1715638115> PMID: 29282314
82. Freeman J, Robinson E, Beckman NG, Bird D, Baggio JA, Anderies JM. The global ecology of human population density and interpreting changes in paleo-population density. *Journal of Archaeological Science*. 2020; 120. <https://doi.org/10.1016/j.jas.2020.105168>
83. Lewis JP, Ryves DB, Rasmussen P, Olsen J, van der Sluis LG, Reimer PJ, et al. Marine resource abundance drove pre-agricultural population increase in Stone Age Scandinavia. *Nature Communications*. 2020; 11(1). <https://doi.org/10.1038/s41467-020-15621-1> PMID: 32332739
84. Loreau M, de Mazancourt C. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology Letters*. 2013; 16:106–15. <https://doi.org/10.1111/ele.12073> PMID: 23346947
85. Sala A, Argueelles CF, Marino ME, Bobillo C, Fenocchio A, Corach D. Genetic Analysis of Six Communities of Mbya-Guarani Inhabiting Northeastern Argentina by Means of Nuclear and Mitochondrial Polymorphic Markers. *Human Biology*. 2010; 82(4):433–56. <https://doi.org/10.3378/027.082.0406> PMID: 21082911
86. Barbieri C, Heggarty P, Yang Yao D, Ferri G, De Fanti S, Sarno S, et al. Between Andes and Amazon: The genetic profile of the Arawak-speaking Yanéscha. *Am J Phys Anthropol*. 2014; 155(4):600–9. <https://doi.org/10.1002/ajpa.22616> PMID: 25229359
87. Crossetti SG, Demarchi DA, Raimann PE, Salzano FM, Hutz MH, Callegari-Jacques SM. Autosomal STR genetic variability in the Gran Chaco native population: Homogeneity or heterogeneity? *American Journal of Human Biology*. 2008; 20(6):704–11. <https://doi.org/10.1002/ajhb.20798> PMID: 18561147

88. Bert F, Corella A, Gené M, Pérez-Pérez A, Turbón D. Mitochondrial DNA diversity in the Llanos de Moxos: Moxo, Movima and Yuracare Amerindian populations from Bolivia lowlands. *Annals of Human Biology*. 2004; 31(1):9–28. <https://doi.org/10.1080/03014460310001616464> PMID: 14742162.
89. Corella A, Bert F, Pérez-Pérez A, Gen M, Turbón D. Mitochondrial DNA diversity of the Amerindian populations living in the Andean Piedmont of Bolivia: Chimane, Mosen, Aymara and Quechua. *Annals of Human Biology*. 2007; 34(1):34–55. <https://doi.org/10.1080/03014460601075819> PMID: 17536754.
90. Casas-Vargas A, Gómez A, Briceño I, Díaz-Matallana M, Bernal JE, Rodríguez JV. High genetic diversity on a sample of pre-Columbian bone remains from Guane territories in northwestern Colombia. *Am J Phys Anthropol*. 2011; 146(4):637–49. <https://doi.org/10.1002/ajpa.21626> PMID: 21990065
91. Di Corcia T, Sanchez Mellado C, Davila Francia TJ, Ferri G, Sarno S, Luiselli D, et al. East of the Andes: The genetic profile of the Peruvian Amazon populations. *Am J Phys Anthropol*. 2017; 163(2):328–38. <https://doi.org/10.1002/ajpa.23209> PMID: 28343372
92. Gnecci-Ruscione GA, Sarno S, De Fanti S, Gianvincenzo L, Giuliani C, Boattini A, et al. Dissecting the Pre-Columbian Genomic Ancestry of Native Americans along the Andes-Amazonia Divide. *Molecular Biology and Evolution*. 2019; 36(6):1254–69. <https://doi.org/10.1093/molbev/msz066> PMID: 30895292
93. Sciar MO, Gouveia MH, Benazzo A, Ghirotto S, Fagundes NJR, Leal TP, et al. Bayesian inferences suggest that Amazon Yunga Natives diverged from Andeans less than 5000 ybp: implications for South American prehistory. *BMC Evolutionary Biology*. 2014; 14. <https://doi.org/10.1186/s12862-014-0174-3> PMID: 25266366
94. Harris DN, Song W, Shetty AC, Levano KS, Cáceres O, Padilla C, et al. Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. *Proceedings of the National Academy of Sciences*. 2018. <https://doi.org/10.1073/pnas.1720798115> PMID: 29946025
95. Hunley K, Healy M. The impact of founder effects, gene flow, and European admixture on native American genetic diversity. *Am J Phys Anthropol*. 2011; 146(4):530–8. <https://doi.org/10.1002/ajpa.21506> PMID: 21913174
96. Hunley K, Gwin K, Liberman B. A Reassessment of the Impact of European Contact on the Structure of Native American Genetic Diversity. *Plos One*. 2016; 11(8). <https://doi.org/10.1371/journal.pone.0161018> PMID: 27579784
97. Yang NN, Mazieres S, Bravi C, Ray N, Wang S, Burley M-W, et al. Contrasting Patterns of Nuclear and mtDNA Diversity in Native American Populations. *Annals of Human Genetics*. 2010; 74:525–38. <https://doi.org/10.1111/j.1469-1809.2010.00608.x> PMID: 20887376
98. Cabana GS, Lewis CM Jr., Tito RY, Covey RA, Cáceres AM, De La Cruz AF, et al. Population Genetic Structure of Traditional Populations in the Peruvian Central Andes and Implications for South American Population History. *Human Biology*. 2014; 86(3):147–65. <https://doi.org/10.13110/humanbiology.86.3.0147> PMID: 25836744